

PH NL  
010450 WO

DOSSIER

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

BEST AVAILABLE COPY

(19) World Intellectual Property Organization  
International Bureau



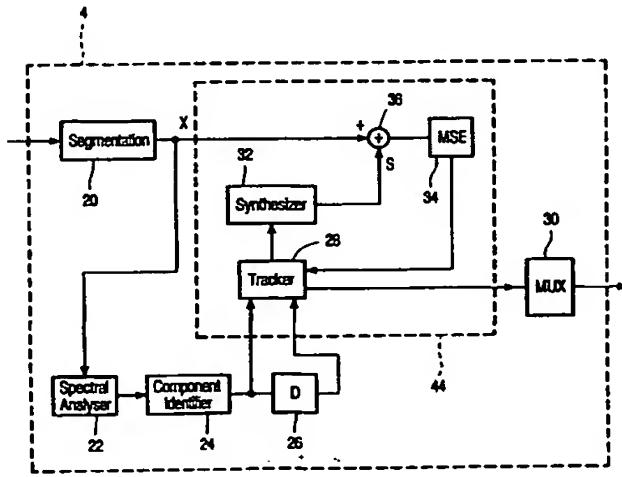
(43) International Publication Date  
28 December 2000 (28.12.2000)

PCT

(10) International Publication Number  
**WO 00/79519 A1**

- (51) International Patent Classification<sup>7</sup>: **G10L 19/02** Albertus, C.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (21) International Application Number: **PCT/EP00/05344** (74) Agent: **GROENENDAAL, Antonius, W., M.**; Internationaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (22) International Filing Date: 8 June 2000 (08.06.2000)
- (25) Filing Language: English (81) Designated States (*national*): CN, JP, KR.
- (26) Publication Language: English (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
- (30) Priority Data:  
99201959.6 18 June 1999 (18.06.1999) EP
- (71) Applicant: **KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]**; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).
- (72) Inventors: **TAORI, Rakesh**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **SLUIJTER, Robert, J.**; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **DEN BRINKER**, Published:  
— With international search report.  
— Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **AUDIO TRANSMISSION SYSTEM HAVING AN IMPROVED ENCODER**



WO 00/79519 A1

(57) Abstract: In a transmission system an audio signal is encoded by an encoder (4), and transmitted by transmit means (6) via a medium (8) to a receiver (10). In the audio encoder (4) audio segments are represented by a plurality of sinusoids each having its own frequency and amplitude. In order to link the audio segments together, it is advantageous to link sinusoids from subsequent segments. The problem in linking the segments is to find which sinusoids should be linked. In the prior art, the components were linked on basis of their frequency differences (closest frequency neighbor criterion). According to the present invention, the linking is done by using an analysis-by-synthesis method. Preferably a preselection is made in order to reduce the complexity of the search. This preselection can be based on the closest frequency neighbor criterion.

Audio transmission system having an improved encoder.

- The present invention relates to a transmission system comprising a transmitter having an audio encoder, said audio encoder comprises segmenting means for dividing an input audio signal into subsequent signal segments, the encoder further comprises track determining means for determining linked signal components present in subsequent signal segments, the transmitter further comprises transmit means for transmitting a signal representing said linked signal components via a transmission medium to a receiver, said receiver comprises receiving means for receiving the signal representing the signal components from the transmission medium, and a decoder for deriving a reconstructed audio signal from said signal representing the signal components.
- 5                  The invention further relates to a transmitter, an encoder, a coding method, a tangible medium, carrying a computer program and a signal, carrying a computer program.

- 10                 A transmission system according to the preamble is known from US 4,885,790.
- 15                 Such transmission systems and audio encoders are used in applications in which audio signals have to be transmitted over a transmission medium with a limited transmission capacity or have to be stored on storage media with a limited storage capacity. Examples of such applications are the transmission of audio signals over the Internet, the transmission of audio signals from a mobile phone to a base station and vice versa and storage of audio signals on a
- 20                 CD-ROM, in a solid state memory or on a hard disk drive.

Different operating principles of audio encoders have been tried to achieve a good audio quality at a modest bit rate. In one of these operating methods, an audio signal to be transmitted is divided into a plurality of segments, having fixed or segment dependent length of 5 - 50 ms. In each of said segments the audio signal is represented by a plurality of signal components, which can be sinusoids that are defined by their amplitudes, their frequencies and possibly their phases.

The transmitting means transmits a representation of the amplitudes and frequencies of the signal components to the receiver. The operations performed by the transmitter can include, channel coding, interleaving and modulation.

The receiving means receive a signal representing the audio signal from a transmission channel and performs operations like demodulation, de-interleaving and channel decoding. The decoder obtains the representation of the audio signal from the receiver and derives a reconstructed audio signal from it by generating a plurality of sinusoids as described 5 by the encoded signal and combining them into an output signal.

In an audio signal often there is a relation between the signal components of the present signal segment and the previous segment. The audio signal can e.g. include a frequency sweep of a sinusoidal component having a duration of a plurality of signal segments. The presence of such kind of signals can be exploited in the encoding process by 10 differentially encoding the frequency and amplitude of signal segments following the first signal segment. This means that instead of the frequency and the amplitude of the signal segment, only the difference between the frequency of the signal component in the present signal segment and in the previous signal component are transmitted. The same can be done with the amplitudes of two subsequent segments.

15 In order to do so, the encoder comprises track determining means for determining which signal components are linked. In the transmission system according to the above mentioned US patent, the frequency difference between signal components in the present frame and in the previous frame are used to determine which frequency components are linked and which are not. However, experiments have shown that this does not always 20 results in an optimum quality of the reconstructed audio signal.

The object of the present invention is to provide a transmission system having an improved quality of the reconstructed audio signal.

25 To achieve said object, the transmission system according to the preamble is characterized in that the tracking means comprises selection means for selecting linked signal components resulting in a minimum error measure between a synthetic audio signal determined on basis of said linked signal components and said input audio signal.

By trying a plurality of candidate linked signal components and selecting that 30 set of linked signal components resulting in a minimum error measure, an improved quality of the reconstructed audio signal is obtained. The error measure can be e.g. the mean squared error or a perceptually weighted error measure. It is observed that it is always necessary to actually generate the synthetic speech signal, but that it is possible to use analytical

methods to determine the optimum parameters of the signal components from the expression of the error signal.

An embodiment of the invention is characterized in that the encoder comprises synthesizing means for deriving the synthetic signal on basis of at least one parameter of said candidate linked signal components in two subsequent signal segments.

By deriving the synthetic signal on basis of at least one parameter (frequency, amplitude) of two subsequent signal segments, the change of said parameters in the synthetic signal is more smooth, which corresponds better to the properties of the input audio signal.

A further embodiment of the invention is characterized in that the synthesizing means are arranged for deriving interpolated signal component parameters from signal component parameters corresponding to at least two subsequent signal segments and in that the synthesizing means are arranged for deriving the synthetic signal on basis of the interpolated signal component parameters.

By using interpolated signal component parameters from signal component parameters of two (or more) subsequent signal segments a naturally sounding synthetic audio signal is obtained.

A still further embodiment of the invention is characterized in that the tracking means comprise pre-selection means for selecting a limited number of candidate linked signal components from all possible linked signal components.

By performing a pre-selection before the actual selection of the linked signal components, it is obtained that the required computational resources are substantially reduced.

A still further embodiment of the invention is characterized in that the pre-selection means are arranged to select as candidate linked signal components signal components having a frequency difference smaller than a predetermined value.

A suitable way of pre-selecting the candidate signal components is the selection of candidate signal components having only a relative small frequency difference. The frequency difference to be allowed is a trade-off between quality and computational complexity. It is however observed that allowing frequency differences above a given value the quality of the reconstructed speech signal hardly improves.

An alternative embodiment of the present invention is characterized in that the tracking means are arranged for extending tracks on the basis of the parameters of linked signal components already determined.

By extending the current track by means of an extrapolation, the complexity of the track determining means is substantially reduced, because the process of trying a large number of candidate signal components can be dispensed with.

5

The present invention will now be explained with reference to the drawings.

Fig. 1 shows a transmission system in which the present invention can be applied.

Fig. 2 shows a first embodiment of an encoder according to the invention for 10 use in the transmission system according to Fig. 1.

Fig. 3 shows a second embodiment of an encoder according to the invention for use in the transmission system according to Fig. 1.

Fig. 4 shows a possible implementation of the tracking block 44 to be used in the encoder according to Fig. 3.

15

In the transmission system according to Fig. 1, the audio signal to be transmitted is applied to an encoder 4 in a transmitter 2. The encoder 4 encodes the input signal as tracks of linked signal components. Said tracks can be represented by a start 20 frequency, a start amplitude and a start phase for the first signal segment, followed by frequency and amplitude differences for the next segments of the track.

A signal representing said tracks is applied to the transmit means 6. The operation of the transmit means 6 can include channel coding, interleaving and modulation. The output of the transmitter 2 is applied to the transmission medium (or recording medium) 8 25 for transmission to the receiver 10. In the receiver 10 the signal from the received signal is applied to the receive means 12. The operations of the receive means 10 can include demodulation, de-interleaving and channel decoding. The output of the receive means 12 is connected to an input of a decoder 14 which is arranged for decoding the encoded audio signal. The operation principle of the decoder 14 is described in the above mentioned US 30 patent.

In the encoder 4 according to Fig. 2, the input signal is applied to segmenting means 20 that divide the input signal in signal segments having fixed or signal dependent segment lengths ranging from 5 – 50 ms. The segments can be partly overlapping, but it is also possible that the segments have no overlap. The segmentation means 20 select a part of the

signal comprising data that is input to the spectral analyzer 22 and data that is input for the closed loop tracking block 44. The appropriate data in either unit is obtained by windowing in the respective units. The segmentation means 20 also generates time markers  $T_k$ , marking the instance in the current signal segment corresponding to the instantaneous tracking parameters  
5 that are to be calculated on basis of the current signal segment,  $T_{k-1}$  corresponds to the instance in the current segment at which the instantaneous tracking parameters have been calculated in the previous iteration. Parameters obtained after closed loop tracking in the  $k^{\text{th}}$  segment are represented using the subscript  $k$  e.g.  $A_k$ . The output of the segmentation means 20 is connected to an input of a spectral analyzer 22 and to a first input of a subtractor 36.

10 The spectral analyzer 22 performs a coarse determination of the frequency spectrum of its input signal. This frequency spectrum forms the basis for determining the linked signal components. The coarse search performed by the spectral analyzer 22 can be based on a moderate size FFT followed by selecting the major peaks in the amplitude spectrum. The output signal of the spectral analyzer 22 is passed to component identification  
15 means 24, which perform a fine search for the parameters of the signal components. This fine search can be done in the neighborhood of the major peaks determined by the spectral analyzer 22 to obtain more accurate estimates of the parameters (amplitude, frequency and phase). Starting values for the amplitude and the frequency for the different signal components can be determined from the amplitude and the frequency of the corresponding peaks in the amplitude  
20 spectrum.

The output signal of the component identification means 24 is passed to a first input of the tracking means 28 and to a delay element 26 which has a delay of one frame period. The output of the delay element 26 is connected to a second input of the tracking means 28. The tracking means 28 starts with determining a limited number of candidate linked  
25 components. The selection of the limited number of candidate linked components can be based on the amplitude and frequency difference. For each signal component of the previous frame, all components in the current frame having a frequency and amplitude difference smaller than predetermined threshold values are added to a table of candidate linked components. Threshold values may be set on the basis of available amplitude and frequency data, and  
30 additionally on the update rate, i.e.  $1/(T_k - T_{k-1})$ . Consequently, a table comprising a plurality of sets each consisting of two linked signal components is obtained. An example of such a table is given below.

	$f_{1,k}=149$	$f_{2,k}=220$	$f_{3,k}=289$	$f_{4,k}=361$	$f_{5,k}=428$	$f_{6,k}=500$	$f_{7,k}=579$
$f_{1,k-1}=131$	x	x					
$f_{2,k-1}=255$	x	x	x	x			
$f_{3,k-1}=380$			x	x	x		
$f_{4,k-1}=495$					x	x	x
$f_{5,k-1}=594$						x	x

The amplitudes and frequencies of the signal components in each set are applied to a synthesizer 32 for deriving a synthetic audio signal from the linked signal components.

5 The synthetic audio signal  $c_{i,j}$ , linking the ith component of Frame k-1 with the jth component of Frame k, can be calculated by using linear interpolation of the frequencies and amplitudes between the boundaries of the frames. The synthetic signal corresponding to one track i,j can be calculated according to:

$$c_{i,j} = M[n] \cdot \cos\left(\frac{\beta}{2}(n - T_{k-1})^2 + \theta_{i,k-1}(n - T_{k-1}) + \phi\right); T_{k-1} \leq n \leq T_k \quad (1)$$

In (1)  $M[n]$  is given by:

$$M[n] = \frac{T_k - n}{T_k - T_{k-1}} A_{i,k-1} + \frac{n - T_{k-1}}{T_k - T_{k-1}} A_{j,k} ; \quad T_{k-1} \leq n \leq T_k \quad (2)$$

10 And  $\beta$  is equal to:

$$\beta = \frac{\theta_{j,k-1} - \theta_{i,k}}{T_k - T_{k-1}} \quad (3)$$

where  $\theta = 2\pi f / fs$ . The phase  $\phi$  can be set to  $\phi_{i,k-1}$ , which is the phase at the end of the previous frame. The phase  $\phi$  can also be derived from a phase value  $\phi_{i,k}$  resulting in a minimum error measure.

An alternative to the synthesizer sketched above can for instance be the well-known overlap add synthesizer.

15 The output of the synthesizer 32 is connected to a second input of the subtractor 36. The subtractor 36 determines the difference between the output signal of the segmentation means 20 and the output signal of the synthesizer 32. The output of the subtractor 36 is connected to an input of a MSE unit 34. This MSE unit 34 determines a mean square value of 20 the output of the subtractor 36 over one frame. Optionally, the MSE unit 34 comprises a weighting filter for perceptually weighting the previously determined mean square value. The output of the MSE unit 34 is connected to a third input of the tracking means 28.

The tracking means 28 use the output signal of the MSE unit 34 to determine which combination of sets of linked signal components results in the smallest output signal of the MSE unit 34. This can be done by performing a full synthesis for all possible combinations of links, but it is also possible to perform a sub-optimal search which requires substantially less computational resources. In this sub-optimal search the synthesis is performed for one component only, and the mean squared error between the input signal and said synthetic component is calculated. This is done for all possible tracks, and a so-called error matrix is constructed. In this matrix all possible tracks are identified together with their corresponding error. Below an example of such a matrix is displayed.

	$f_{1,k-l}=149$	$f_{2,k-l}=220$	$f_{3,k-l}=289$	$f_{4,k-l}=361$	$f_{5,k-l}=428$	$f_{6,k-l}=500$	$f_{7,k-l}=579$
$f_{1,k-l}=131$	52876	62929					
$f_{2,k-l}=255$	76778	90541	55049	69752			
$f_{3,k-l}=380$			80849	89816	63016		
$f_{4,k-l}=495$					70948	74605	69535
$f_{5,k-l}=594$						73881	70260

10

In the matrix only the errors are included for the components corresponding to the preselected links.

Although in the example shown above all links are shown, it is possible to leave out the links which cause an increase of the error instead of a decrease. When the possible links are resolved, it must be the case that a signal component in the current frame has only one single corresponding component in the previous frame, and that a signal component in the previous frame has only one single corresponding component in the current frame. In the above matrix this means that rows and columns may not have multiple entries.

One possible sub-optimal way of finding the final is to first select the element in each column having the smallest error, and subsequently selecting the element in each row having the smallest error. After selecting the single elements in the columns, the matrix changes into:

	$f_{1,k}=149$	$f_{2,k}=220$	$f_{3,k}=289$	$f_{4,k}=361$	$f_{5,k}=428$	$f_{6,k}=500$	$f_{7,k}=579$
$f_{1,k-1}=131$	52876	62929					
$f_{2,k-1}=255$			55049	69752			
$f_{3,k-1}=380$					63016		
$f_{4,k-1}=495$							69535
$f_{5,k-1}=594$						73881	

After subsequently selecting the single elements in the columns, the following final link matrix is obtained.

	$f_{1,k}=149$	$f_{2,k}=220$	$f_{3,k}=289$	$f_{4,k}=361$	$f_{5,k}=428$	$f_{6,k}=500$	$f_{7,k}=579$
$f_{1,k-1}=131$	52876						
$f_{2,k-1}=255$			55049				
$f_{3,k-1}=380$					63016		
$f_{4,k-1}=495$							69535
$f_{5,k-1}=594$						73881	

5

It is also possible to start with selecting the element in each row having the smallest error and subsequently selecting the element in each column having the smallest error. It is also possible that both possibilities are tried, and the results of the selection method leading to the smallest mean squared error (= the sum of all matrix elements) are chosen.

10 At the end of the procedure the tracks to be continued are available. The signal components in the previous frames that are not linked are either the end of a track or isolated points. The components in the current frame that could not be linked are either a newly appearing track or an isolated point. The inclusion of the parameters of the isolated points in the signal to be transmitted can be based on subjective measures such as masking threshold or  
15 on objective measures such as reduction in the error or a combination of both. Since inclusion of these isolated components also has an impact on the bitrate, the bit budget can also be considered in the process.

20 After the selection means have selected the linked signal components and the isolated points to be transmitted, signals representing the tracks of linked components are applied to a multiplexer 30 which combines them into a multiplex signal which is suitable for transmission.

In the encoder 4 according to Fig. 3, the segmentation unit 20, the function of which has already been described, selects a part of the signal comprising data that is input to the spectral analyzer 22 and data that is input for the closed loop tracking block 44. The appropriate data in either unit is obtained by windowing in the respective units.

5 A memory unit 32 holds a description relevant information from the previous segment  $k-1$ . The control unit 40 uses the input signal received from the memory unit 32 and possibly the input from the spectral analyzer 22 to make an estimate of track continuation parameters. The continuation data comprises information about which tracks will be continued and the ones that will be discontinued. For each track that will be continued, at least one  
10 estimate of the frequency location, to which it will be continued is provided.

The output of the control unit 40 along with the current input segment serve as the inputs to the closed loop tracking block 44. The output of the closed loop tracking block 44 consists of track data (indicated using TD in Figure 3) of both the continuing tracks and the newly appearing tracks. The output FD of tracking block 44 in Figure 3 contains auxiliary  
15 information in addition to the track data TD. The auxiliary information contains information that is necessary to perform local synthesis and could contain additional information such as the reduction in error that was obtained.

The track data TD is fed to the multiplexer 30, which combines the information in to a multiplex signal which is suitable for transmission.

20 One of the tasks of the control unit 40 is to decide whether a track should be continued in the current segment. This decision is taken on the basis of the available track information thus far including the auxiliary information. The other task is to compute one or more initial estimates of the frequency at  $T_k$  for each track that should be continued. This computation can be carried out for varying levels of sophistication, starting from zero<sup>th</sup> order  
25 extrapolation where only frequency information at  $T_{k-1}$  is required, to higher-order prediction using frequency information at  $T_i$  with  $i < k-1$ . In this computation, use can be made from the additional information available at the output of the spectral analyzer 22.

A possible implementation of the tracking block 44 is shown in Fig. 4.

The outputs of the control unit 40 and the segmentation unit 20 are input to the continuation  
30 tracker 50. The continuation tracker implements a search procedure so as to deliver the optimum parameters corresponding to continuing tracks and the auxiliary information. Using this, a signal s can be synthesized by the synthesizer 52. The output of the synthesizer 52 along with the output of the segmentation unit 20 and possibly the output of the control unit 40 is used by the tracker of newly appearing components 54 to deliver the optimum parameters for

the newly appearing tracks along with the auxiliary information. The outputs of the trackers 50 and 54 are used by the splitter 56 to deliver track data (TD) and feedback data (FD).

5 The continuation tracker 50 implements a search procedure to deliver an improved estimate of the frequency  $\theta_m$  at time  $T_k$ , along with the optimum amplitude  $A_m$  and phase  $\phi_m$  and auxiliary information. The search itself is based upon minimizing the weighted squared error, which can be expressed as:

$$E = \sum_n w_k(n) \cdot [x(n) - s(n)]^2 \quad (4)$$

where  $w_k$  is the weighting function in the  $k^{\text{th}}$  segment,  $x[n]$  is the input segment,  $s[n]$  is the synthesized segment and  $n$  addresses the interval defined by the weighting function. The weighting function is generated on the basis of segment length and the markers  $T_k - T_{k-1}$ . Let

10  $s(n)$  represent the synthesized signal according to:

$$s(n) = \sum_{m=1}^v a_m(n) \cos(\phi_m(n)) + b_m(n) \sin(\phi_m(n)) \quad (5)$$

In (5)  $v$  is the number of continuing tracks in the current frame and the coefficients  $a_m(n)$ ,  $b_m(n)$  and  $\phi_m(n)$  in (7) are calculated according to:

$$a_m(n) = a_m(T_{k-1}) + \{a_m(T_k) - a_m(T_{k-1})\} \frac{n - T_{k-1}}{T_k - T_{k-1}} \quad (6)$$

$$b_m(n) = b_m(T_k) \frac{n - T_{k-1}}{T_k - T_{k-1}} \quad (7)$$

$$\theta_m(n) = \theta_m(T_{k-1}) + \{\theta_m(T_k) - \theta_m(T_{k-1})\} \frac{n - T_{k-1}}{T_k - T_{k-1}} \quad (8)$$

$$\phi_m(n) = \phi_m(T_{k-1}) + \sum_{l=T_{k-1}}^{n-1} \theta(l) \quad (9)$$

$$\phi_m(T_{k-1}) = \phi_m(T_{k-1}) \quad (10)$$

$$a_m(T_{k-1}) = A_m(T_{k-1}) \quad (11)$$

15 Thus the minimization of (6) can be expressed as:

$$\min E \\ \theta_m(T_k), a_m(T_k), b_m(T_k) \quad (12)$$

where the minimum is taken over all  $m$  according to  $1 \leq m \leq v(T_{k-1})$ . It is noted, however, that  $a_m(T_k)$  and  $b_m(T_k)$  appear linearly in  $s(n)$ . This implies that their optimal values according to the given minimization criterion are the solution of a set of normal equations, thereby simplifying the search procedure and reducing the problem to minimization over  $\theta_m(T_k)$ .

5

By using (7) to (12), (4) can be rewritten as:

$$E = \sum_n w_k(n) \left[ x(n) - \sum_{m=1}^{v(T_{k-1})} a_m(T_{k-1}) \left( 1 - \frac{n - T_{k-1}}{T_k - T_{k-1}} \right) \cos(\phi_m(n)) \right. \\ \left. - \sum_{m=1}^{v(T_{k-1})} a_m(T_k) \left( \frac{n - T_{k-1}}{T_k - T_{k-1}} \right) \cos(\phi_m(n)) - \sum_{m=1}^{v(T_{k-1})} b_m(T_k) \left( \frac{n - T_{k-1}}{T_k - T_{k-1}} \right) \sin(\phi_m(n)) \right]^2 \quad (13)$$

Each iteration of the search procedure is started by selecting values of  $\theta_m(T_k)$  and calculating the optimum corresponding values of  $a_m(T_k)$  and  $b_m(T_k)$ . In order to calculate the values of  $a_m(T_k)$  and  $b_m(T_k)$ , the following substitutions are performed:

$$\tilde{x}(n) = x(n) - \sum_{m=1}^{v(T_{k-1})} a_m(T_{k-1}) \left( 1 - \frac{n - T_{k-1}}{T_k - T_{k-1}} \right) \cos(\phi_m(n)) \quad (14)$$

$$p_{2m-1}(n) = \frac{n - T_{k-1}}{T_k - T_{k-1}} \cos(\phi_m(n)) \quad (15)$$

$$p_{2m}(n) = \frac{n - T_{k-1}}{T_k - T_{k-1}} \sin(\phi_m(n)) \quad (16)$$

$$c_{2m-1} = a_m(T_k) \quad (17)$$

$$c_{2m} = b_m(T_k) \quad (18)$$

Substituting (15) to (19) into (14) gives the following criterion to be minimized:

$$E = \sum_n w_k(n) \left[ \tilde{x}(n) - \sum_{m=1}^{2v(T_{k-1})} c_m \cdot p_m(n) \right]^2 \quad (19)$$

10

The minimum of  $E$  is found by differentiating (19) with respect to each of the values  $c_m$  and setting each derivative to zero. This leads to the following set of equations from which the constants  $c_1, c_2, \dots, c_{2v-1}, c_{2v}$  have to be determined.

$$\begin{aligned}
 \sum_n w_k(n) \cdot \tilde{x}(n) \cdot p_1(n) - \sum_n w_k(n) \cdot \sum_{m=1}^{2v(T_{k-1})} c_m \cdot p_m(n) \cdot p_1(n) &= 0 \\
 \sum_n w_k(n) \cdot \tilde{x}(n) \cdot p_2(n) - \sum_n w_k(n) \cdot \sum_{m=1}^{2v(T_{k-1})} c_m \cdot p_m(n) \cdot p_2(n) &= 0 \\
 \vdots &\quad \vdots \\
 \sum_n w_k(n) \cdot \tilde{x}(n) \cdot p_{2v-1}(n) - \sum_n w_k(n) \cdot \sum_{m=1}^{2v(T_{k-1})} c_m \cdot p_m(n) \cdot p_{2v-1}(n) &= 0 \\
 \sum_n w_k(n) \cdot \tilde{x}(n) \cdot p_{2v}(n) - \sum_n w_k(n) \cdot \sum_{m=1}^{2v(T_{k-1})} c_m \cdot p_m(n) \cdot p_{2v}(n) &= 0
 \end{aligned} \tag{20}$$

(21) can be cast in matrix notation according to  $\bar{d} - A\bar{c} = \bar{0}$  with  $\bar{d}$ ,  $A$  and  $\bar{c}$  being equal to:

$$\bar{d} = \left( \begin{array}{c} \sum_n w(n) \cdot \tilde{x}(n) \cdot p_1(n) \\ \sum_{n_1} w(n) \cdot \tilde{x}(n) \cdot p_2(n) \\ \vdots \\ \sum_n w(n) \cdot \tilde{x}(n) \cdot p_{2v-1}(n) \\ \sum_n w(n) \cdot \tilde{x}(n) \cdot p_{2v}(n) \end{array} \right) \tag{21}$$

$A =$

$$\left( \begin{array}{cccc} \sum_n w_k(n) p_1(n) p_1(n) & \sum_n w_k(n) p_1(n) p_2(n) & \dots & \sum_n w_k(n) p_1(n) p_{2v-1}(n) & \sum_n w_k(n) p_1(n) p_{2v}(n) \\ \sum_n w_k(n) p_2(n) p_1(n) & \sum_n w_k(n) p_2(n) p_2(n) & \dots & \sum_n w_k(n) p_2(n) p_{2v-1}(n) & \sum_n w_k(n) p_2(n) p_{2v}(n) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sum_n w_k(n) p_{2v-1}(n) p_1(n) & \sum_n w_k(n) p_{2v-1}(n) p_2(n) & \dots & \sum_n w_k(n) p_{2v-1}(n) p_{2v-1}(n) & \sum_n w_k(n) p_{2v-1}(n) p_{2v}(n) \\ \sum_n w_k(n) p_{2v}(n) p_1(n) & \sum_n w_k(n) p_{2v}(n) p_2(n) & \dots & \sum_n w_k(n) p_{2v}(n) p_{2v-1}(n) & \sum_n w_k(n) p_{2v}(n) p_{2v}(n) \end{array} \right) \tag{22}$$

$$\bar{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{2v-1} \\ c_{2v} \end{pmatrix} \quad (23)$$

The value of  $\bar{c}$  can now be calculated according to:

$$\bar{c} = A^{-1} \cdot \bar{d} \quad (24)$$

The optimization can be carried out over all tracks simultaneously or track by track. In the latter case, the optimization block decides on the ordering. Such a decision can for instance be made on the basis of the previously strength expressed in  $A_m$ . Conventional techniques to terminate the search procedure can be invoked. From  $a_m$ ,  $b_m$  and  $\phi_m$  at  $T_k$  the values  $A_m$  and

$\phi_m$  at  $T_k$  are calculated according to:

$$\begin{aligned} A_m &= \sqrt{a_m^2 + b_m^2} \\ \phi_m &= \phi_m - \Psi_m \end{aligned} \quad (25)$$

Where  $\psi_m$  corresponds to the angle of the complex number  $a_m + jb_m$ .

The tracker of newly appearing components 54 is similar in functionality to the continuation tracker 50. The differences occur due to the fact that the tracker of newly appearing components 54 identifies newly appearing tracks rather than continuing tracks.

Accordingly, the target signal against which the error is minimized may be derived from the segment synthesized thus far,  $s[n]$  together with the input segment  $x[n]$ . In its simplest form this may be derived by subtracting  $s$  from  $x$ . Additionally, the amplitude and the phase values  $A_m$  and  $\phi_m$  at  $T_k$  are set to 0. During a newly appearing track one has to make an assumption on the track of the frequency or the phase within the current segment. That the frequency is held constant during the birth is a suitable assumption. Thus, in its simplest form, the frequency value  $\theta_m$  at  $T_{k-1}$  may be set in accordance with a constant frequency track.

It is observed that the present invention can be implemented in dedicated hardware, in software running on a DSP or on a general purpose computer. The present invention can be embodied in a tangible medium such as a CD-ROM or DVD-ROM carrying a

computer program for executing an encoding method according to the invention. The invention can also be embodied as a signal transmitted over a data network such as the Internet, or a signal transmitted by a broadcast service.

## CLAIMS:

1. Transmission system comprising a transmitter having an audio encoder, said audio encoder comprises segmenting means for dividing an input audio signal into subsequent signal segments, the encoder further comprises track determining means for determining linked signal components present in subsequent signal segments, the transmitter further comprises transmit means for transmitting a signal representing said linked signal components via a transmission medium to a receiver, said receiver comprises receiving means for receiving the signal representing the signal components from the transmission medium, and a decoder for deriving a reconstructed audio signal from said signal representing the signal components, characterized in that the tracking means comprises selection means for selecting linked signal components resulting in a minimum error measure between a synthetic audio signal determined on basis of said linked signal components and said input audio signal.
2. Transmission system according to claim 1, characterized in that the encoder comprises synthesizing means for deriving the synthetic signal on basis of at least one parameter of said candidate linked signal components in two subsequent signal segments.
3. Transmission system according to claim 2, characterized in that the synthesizing means are arranged for deriving interpolated signal component parameters from signal component parameters corresponding to at least two subsequent signal segments and in that the synthesizing means are arranged for deriving the synthetic signal on basis of the interpolated signal component parameters.
4. Transmission system according to claim 1, 2 or 3, characterized in that the tracking means comprise pre-selection means for selecting a limited number of candidate linked signal components from all possible linked signal components.
5. Transmission system according to claim 4, characterized in that the pre-selection means are arranged to select as candidate linked signal components signal components having a frequency difference smaller than a predetermined value.

6. Transmission system according to claim 1, characterized in that the tracking means are arranged for extending tracks on basis of the parameters of linked signal components already determined.

5 7. Transmission system according to claim 6, characterized in that the tracking means are arranged for determining linked signal components on basis of a synthesis of one track.

8. Transmitter having an audio encoder, said audio encoder comprises segmenting means for dividing an input audio signal into subsequent signal segments, the encoder further comprises track determining means for determining linked signal components present in subsequent signal segments, the transmitter further comprises transmit means for transmitting a signal representing said linked signal components, characterized in that the tracking means comprises selection means for selecting linked signal components resulting in a minimum error measure between a synthetic audio signal determined on basis of said linked signal components and said input audio signal.

9. Transmitter according to claim 8, characterized in that the encoder comprises synthesizing means for deriving the synthetic signal on basis of at least one parameter of said candidate linked signal components in two subsequent signal segments.

10. Transmitter according to claim 9, characterized in that the synthesizing means comprise weighting means for weighting the signal component in the two subsequent signal segments with a corresponding weighting function, and combining means for combining the weighted signal segments into the synthetic audio signal.

11. Transmitter according to claim 8, 9 or 10, characterized in that the tracking means comprise pre-selection means for selecting a limited number of candidate linked signal components from all possible linked signal components.

30 12. Transmitter according to claim 8, characterized in that the tracking means are arranged for extending tracks on basis of the parameters of linked signal components already determined.

13. Audio encoder comprising segmenting means for dividing an input audio signal into subsequent signal segments, the encoder further comprises track determining means for determining linked signal components present in subsequent signal segments, the transmitter further comprises transmit means for transmitting a signal representing said linked signal components, characterized in that the tracking means comprises selection means for selecting linked signal components resulting in a minimum error measure between a synthetic audio signal determined on basis of said linked signal components and said input audio signal.

5            14. Coding method comprising dividing an input audio signal into subsequent signal segments, determining linked signal components present in subsequent signal segments, transmitting a signal representing said linked signal components, characterized in that the method comprises selecting linked signal components resulting in a minimum error measure between a synthetic audio signal determined on basis of said linked signal components and said input audio signal.

10            15. Tangible medium comprising a computer program for performing the coding method according to claim 13.

15            16. Signal carrying a computer program for performing the coding method according to claim 13.

1/2

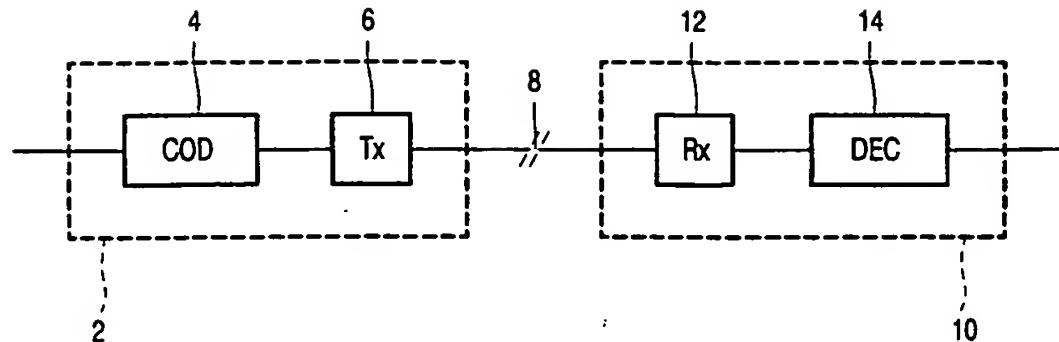


FIG. 1

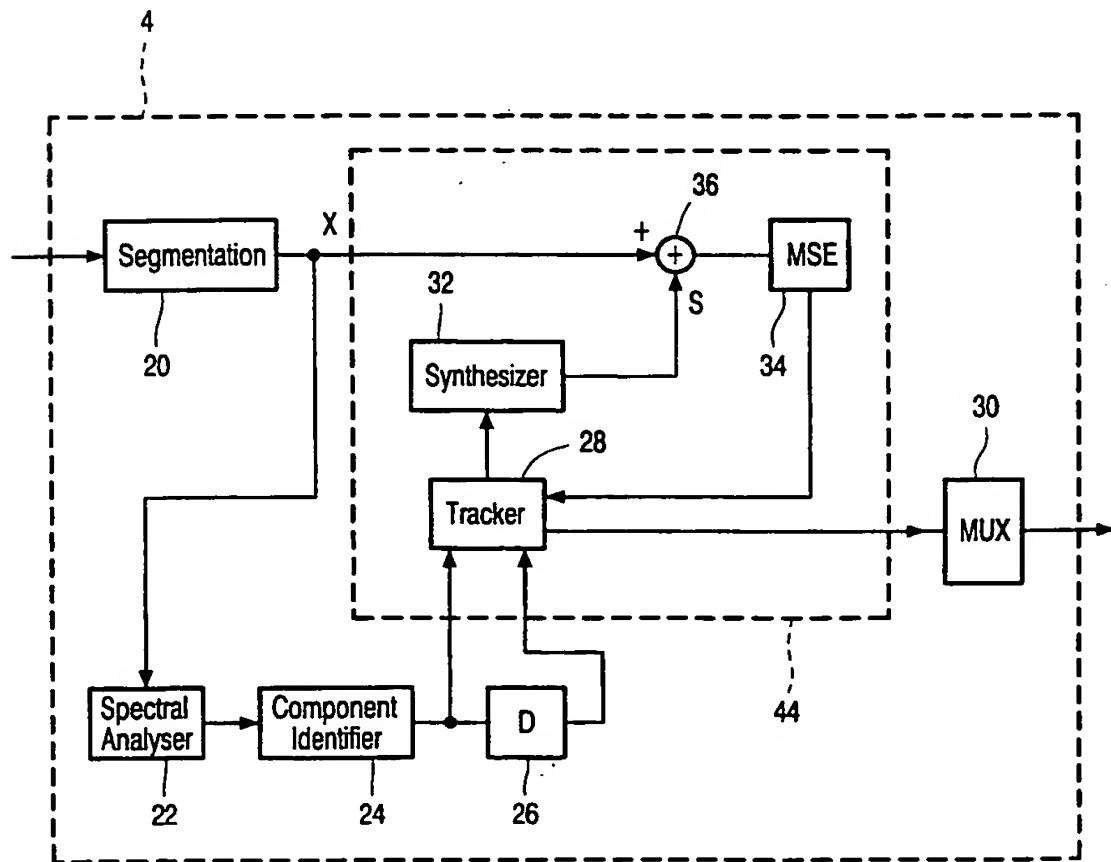


FIG. 2

2/2

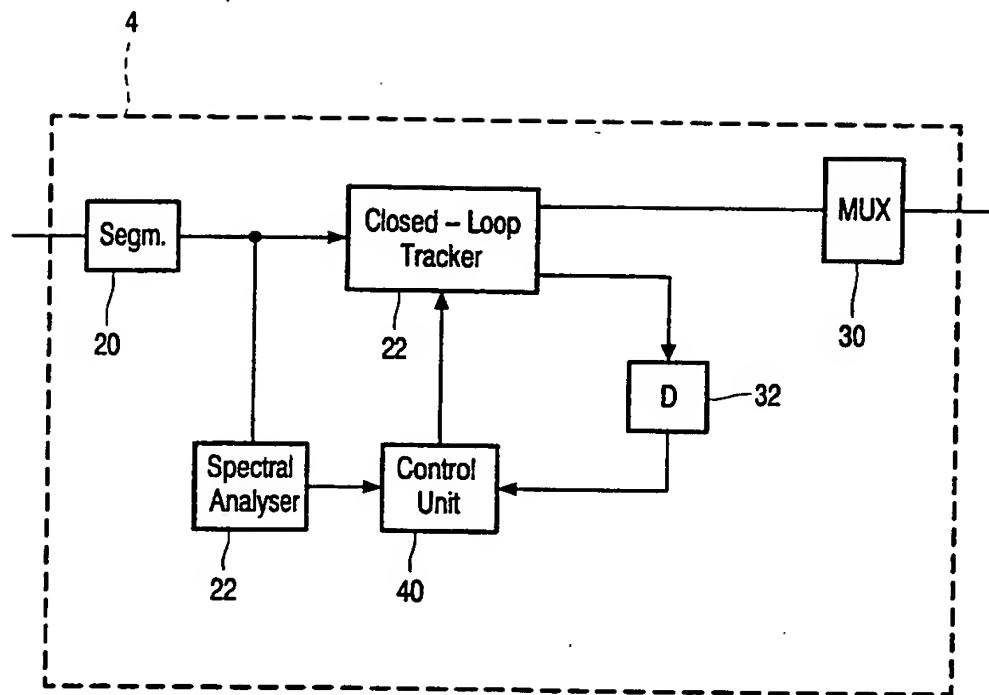


FIG. 3

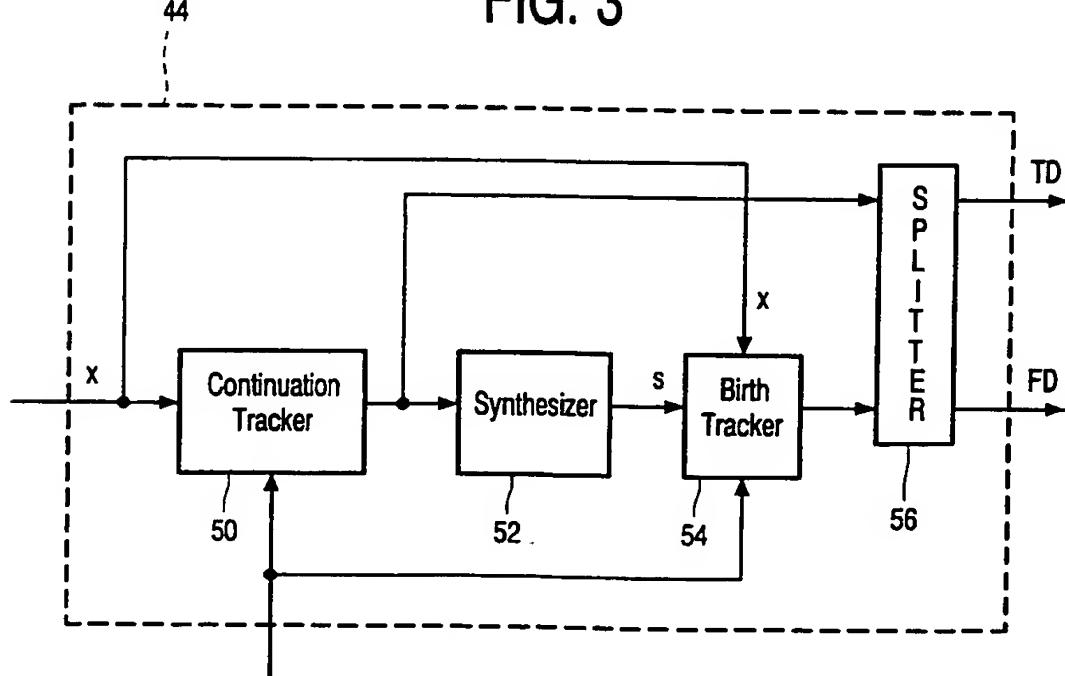


FIG. 4

# INTERNATIONAL SEARCH REPORT

1. International Application No  
PCT/EP 00/05344

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G10L19/02

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, PAJ, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 89 09985 A (MASSACHUSETTS INST TECHNOLOGY) 19 October 1989 (1989-10-19) page 3, paragraph 3 page 13 page 15 ---	1,8,13, 14
Y	US 4 885 790 A (MCAULAY ROBERT J ET AL) 5 December 1989 (1989-12-05) cited in the application column 2, line 3 - line 13 column 2, line 34 - line 65 column 6, line 24 -column 7, line 23 ---	1,8,13, 14

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

\* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "G" document member of the same patent family

Date of the actual completion of the international search

9 October 2000

Date of mailing of the international search report

26/10/2000

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.  
Fax: (+31-70) 340-3016

Authorized officer

Ramos Sánchez, U

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 00/05344

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5 504 833 A (GEORGE E BRYAN ET AL) 2 April 1996 (1996-04-02) column 2, line 51 - line 60 column 7, line 27 - line 39 column 9, line 33 - line 67 column 21, line 44 - line 51 --- 	1,8,13, 14
A	EP 0 337 636 A (AMERICAN TELEPHONE & TELEGRAPH) 18 October 1989 (1989-10-18) page 3, line 44 - line 55 page 6, line 19 - line 51 --- 	1,8,13, 14
P,X	TAORI R ET AL: "Closed-loop tracking of sinusoids for speech and audio coding" IEEE WORKSHOP ON SPEECH CODING PROCEEDINGS. MODEL, CODERS, AND ERROR CRITERIA, PORVOO, FINLAND, 20 - 23 June 1999, pages 1-3, XP002149588 IEEE, Piscataway, NJ, USA ISBN: 0-7803-5651-9 the whole document --- 	1-16
P,A	PURNHAGEN H: "Advances in parametric audio coding" PROCEEDINGS OF THE 1999 IEEE WORKSHOP ON APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS (WASPAA'99), NEW PALTZ, NY, USA , 17 - 20 October 1999, pages 31-34, XP002149587 IEEE, Piscataway, NJ, USA ISBN: 0-7803-5612-8 page 32, left-hand column, paragraph 2 - paragraph 3 ----- 	1-16

# INTERNATIONAL SEARCH REPORT

Information on patent family members

b. International Application No

PCT/EP 00/05344

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
WO 8909985	A 19-10-1989	US 4937873 A			26-06-1990
		AU 3736289 A			03-11-1989
		CA 1337665 A			28-11-1995
US 4885790	A 05-12-1989	US RE36478 E			28-12-1999
		AU 597573 B			07-06-1990
		AU 5620886 A			13-10-1986
		CA 1243122 A			11-10-1988
		EP 0215915 A			01-04-1987
		JP 2759646 B			28-05-1998
		JP 62502572 T			01-10-1987
		US 4856068 A			08-08-1989
		WO 8605617 A			25-09-1986
		US 4937873 A			26-06-1990
US 5504833	A 02-04-1996	US 5327518 A			05-07-1994
		EP 0759201 A			26-02-1997
		WO 9530983 A			16-11-1995
		WO 9304467 A			04-03-1993
EP 0337636	A 18-10-1989	US 5179626 A			12-01-1993
		CA 1336456 A			25-07-1995
		DE 68916831 D			25-08-1994
		JP 2203398 A			13-08-1990

**THIS PAGE BLANK (USPTO)**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)